PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04N 7/24, 7/26, 7/173

A1

(11) International Publication Number: WO 99/00984

(43) International Publication Date: 7 January 1999 (07.01.99)

(21) International Application Number: PCT/IL98/00300

(22) International Filing Date: 25 June 1998 (25.06.98)

(30) Priority Data:
08/884,531 26 June 1997 (26.06.97) US

(71) Applicant (*for all designated States except US*): CITRIX SYSTEMS, INC. [US/US]; 6400 N.W. 6th Way, Fort Lauderdale, FL 33309 (US).

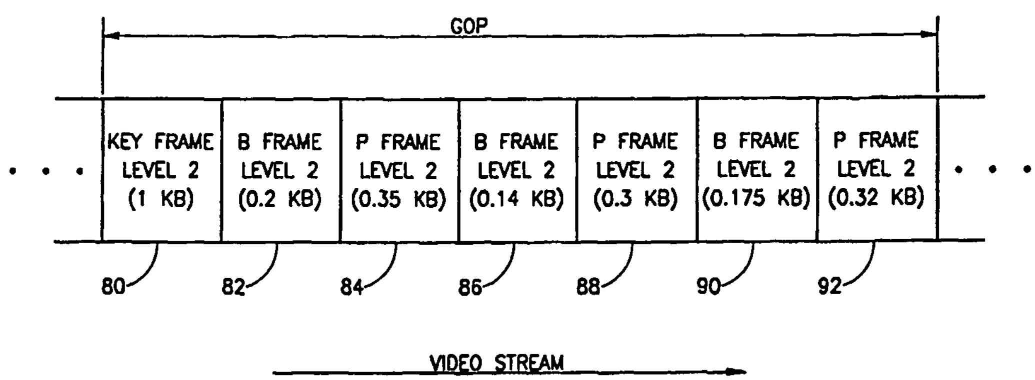
(72) Inventors; and
(75) Inventors/Applicants (*for US only*): AHARONI, Amir [IL/IL]; Portzei Haderch Street 79, 46274 Herzelia (IL). KHIRMAN, Stas [IL/IL]; Rashi Street 11/7, 45287 Hod Hasharon (IL). TAITS, Eugene [IL/IL]; Hertzel Street 12, 44444 Kfar Saba (IL). ARIEL, Oren [IL/IL]; Rashi Street 7, 47207 Ramat Hasharon (IL).

(74) Agents: BAUER, Steven, M. et al.; Testa, Hurwitz & Thibeault, LLP, High Street Tower, 125 High Street, Boston, MA 02110 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: SYSTEM FOR ADAPTIVE VIDEO/AUDIO TRANSPORT OVER A NETWORK



(57) Abstract

A system for adaptively transporting video over networks wherein the available bandwidth varies with time. The system comprises a video/audio codec that functions to compress, code, decode and decompress video streams that are transmitted over networks having available banwidths that vary with time and location. Depending on the channel bandwidth, the system adjusts the compression ratio to accomodate a plurality of bandwidths ranging from 20 Kbps for POTS to several Mbps for switched LAN and ATM environments. Bandwidth adjustability is provided by offering a trade off between video resolution, frame rate and individual frame quality. The system generates a video data stream comprised of Key, P and B frames from a raw source of video. Each frame type is further comprised of multiple levels of data representing varying degrees of quality. In addition, several video server platforms can be utilized in tandem to transmit video/audio information with each video server platform transmitting information for a single compression/resolution level.

SUMMARY OF THE INVENTION

The present invention is a system for adaptively transporting video over networks wherein the available bandwidth varies with time. The present invention has application to any type of network including those that utilize the Internet Protocol (IP) such as the Internet or other TCP/IP networks. The system comprises a video/ audio codec or coder/decoder that functions to compress, code, decode and decompress video streams that are transmitted over networks having available bandwidths that vary with time and location. Depending on the channel bandwidth, the system adjusts the compression ratio to accommodate a plurality of bandwidths ranging from 20 Kbps for plain old telephone service (POTS) to several Mbps for switched LAN and ATM environments. Bandwidth adjustability is provided by offering a trade off between video resolution (e.g., 160 x 120, 320 x 240, 640 x 480), frame rate (e.g., 30 fps, 15 fps, 7.5 fps) and individual frame quality. This flexibility is useful for different applications that stress different requirements.

The system functions to generate a prioritized video data stream comprising multiple levels from a raw source of video. This video stream is stored in a file and accessed by the video server when servicing clients. In operation, the video client only receives a subset of the levels. The levels are chosen to have a suitable data content to match that of the network connection. This permits a better fit between network bandwidth consumed and video image quality. Each of the levels is built on top of the previous levels, with the higher levels providing incremental information not present in the lower levels. This ensures that bandwidth is not wasted on the client end or on the encoder/server side. The system generates the video stream that is sent to the client such that a loss of any individual packet on the network will not cause sustained degraded quality at the client.

The scaleable compression performed by the system is suitable for transparent video within an Internet environment characterized by large diversity and heterogeneity. The system functions to match the image quality of the video data being transported with the wide variations in available network bandwidth. In

addition, the system can adjust the video data to match the differences in available computing power on the client computer system. The system, utilizing 'best effort' protocols such as those found on the Internet, adapts to the time varying nature of the available bandwidth.

There is therefore provided in accordance with the present invention a method of transporting video over a network channel, comprising the steps of compressing a raw video source into a plurality of frames, each frame comprising a plurality of levels, each level corresponding to a particular degree of compression, estimating the bandwidth of the network channel, selecting one of the plurality of levels of each frame to transmit over the network channel in accordance with the bandwidth estimate whereby the level selected optimizes the use of the bandwidth of the network channel, and sending the selected level of each frame over the network channel.

The step of compressing comprises the step of compressing the raw video source into a plurality of different types of frames, each frame type containing different amount of video content information, the plurality of different types of frames grouped so as to form a video stream consisting of a plurality of group of pictures (GOP) sequences. The step of compressing comprises the step of compressing the raw video source into Key, P and B type frames, the Key, P and B frames generated so as to form a video stream consisting of a plurality of group of pictures (GOP) sequences.

There is also provided in accordance with the present invention a method of transporting video from a video server to a video client over a network channel, comprising the steps of compressing data from a raw video source so as to generate a plurality of frames, each frame being of a particular frame type, each frame type containing a particular amount of video content information, each frame comprising a plurality of levels, each level corresponding to a particular degree of compression, estimating the bandwidth of the network channel, determining the amount of video information waiting to be displayed at the video client, selecting one of the plurality of levels of each frame to send over the network channel in accordance with the bandwidth estimate whereby the level

selected optimizes the use of the bandwidth of the network channel, choosing which frames having a particular frame type to send over the network channel in accordance with the amount of video information waiting to be displayed at the video client, and sending the chosen frames having a particular frame type and of the selected level over the network channel.

Further, there is provided in accordance with the present invention a video server for transporting video from a video source over a network channel to a video client, the video source consisting of a plurality of frames of video data, each frame of video data consisting of multiple compression levels and being of a particular type, the video server comprising receiver means for inputting frames of video data from the video source, sending means coupled to the receiver means, the sending means for determining which compression level within the frame and which frames having a particular type to transmit in accordance with the estimated available bandwidth of the network channel, the sending means for encapsulating the frames of video data into a plurality of packets for transmission over the network channel, and a controller for managing the operation of the receiver means and the sending means whereby the rate of transmission of the sending means is maintained so as to match the available bandwidth of the network channel.

In addition, the sending means comprises a rate control unit for measuring the available bandwidth of the network channel, a frame selector for inputting video frame data output by the receiver means, the frame selector outputting frames of a particular compression level in accordance with the bandwidth measured by the rate control unit, a packet generator for inputting video frame data output by the frame selector, the packet generator for encapsulating the video frame data into a plurality of packets for transmission, the packet generator determining which frames having a particular type are to be transmitted, a packet transmitter for placing onto the network channel the plurality of packets output by the packet generator, and a receiver for receiving acknowledgments sent by the video client over the network channel in response to packets received thereby.

There is further provided in accordance with the present invention a method of measuring the bandwidth of a network channel connecting a sender to a receiver, the method comprising the steps of the sender transmitting a plurality of packets to the receiver over the network channel to yield a particular number of bytes online, the receiver transmitting to the sender acknowledgments in response to the receipt of the packets by the receiver, measuring the reception bandwidth of the packets by the receiver, increasing the number of bytes online until the rate of increase of the reception bandwidth decreases to within a predetermined threshold, and estimating the bandwidth of the network channel to be the reception bandwidth at the receiver.

In addition, there is provided in accordance with the present invention a method of maintaining a maximum number of bytes online in a network channel connecting a sender to a receiver, the network channel having a particular available bandwidth, the method comprising the steps of determining the number of bytes sent (BytesSent) to the receiver utilizing sender related data concerning the previous packet sent and the last packet sent, determining the number of bytes received (BytesRec) by the receiver utilizing receiver related data concerning the previous packet received and the last packet received, calculating the sending rate (SendRate) in accordance with the following equation $\mathrm{SendRate} = \dfrac{\mathrm{BytesSent}}{\mathrm{TimeToSend(PreviousResp)} - \mathrm{TimeToSend(LastResp)}}$, calculating the receiving rate (RecRate) in accordance with the following equation $\mathrm{RecRate} = \dfrac{\mathrm{BytesRec}}{\mathrm{TimeToRec(PreviousResp)} - \mathrm{TimeToRec(LastResp)}}$, comparing the sending rate to the receiving rate, increasing the sending rate if the sending rate is less than or equal to the receiving rate, and decreasing the sending rate if the sending rate is greater than the receiving rate.

There is also provided in accordance with the present invention a method of transporting video from a video server to a video client over a network channel, comprising the steps of compressing data from a raw video source so as to generate a plurality of frames, each frame being of a particular frame type,

compress, code, decode and decompress video streams that are transmitted over the network 20 into a compressed video and audio file. The compressed file may be in any suitable format such as Audio Video Interleaved (AVI) format. Note that the network may comprise any type of network, TCP/IP or otherwise including the Internet. The generation of the compressed video and audio file 16 can be performed either online or off-line. Typically, the video and audio file is generated off-line. Note that any suitable method of video compression can be utilized in the present invention such as described in connection with the Motion Pictures Expert Group (MPEG)-1, MPEG-2 or MPEG-4 standards.

One important aspect of the invention is that although the available bandwidth of the network may vary with time and location, the quality of the transmitted video is varied in accordance with the available bandwidth. Depending on the channel bandwidth, the system adjusts the compression ratio to accommodate a plurality of bandwidths ranging from 20 Kbps for plain old telephone service (POTS) to several Mbps for switched LAN environments. Bandwidth adjustability is provided by offering a trade off between video resolution (e.g., 160 x 120, 320 x 240, 640 x 480), frame rate (e.g., 30 fps, 15 fps, 7.5 fps) and individual frame quality. This flexibility is useful for different applications that stress different requirements.

The system functions to generate a prioritized video data stream comprising multiple levels from a raw source of video 12. This video stream is stored in a file (compressed video and audio file 16 in Figure 1) and accessed by the video server 18 when servicing clients 22. In operation, the video client only receives a subset of the levels that form the video and audio file 16. The levels are chosen to have a suitable data content to match that of the network connection between server and client. This permits a better fit between network bandwidth consumed and video image quality. Each of the levels is built on top of the previous levels, with the higher levels providing incremental information not present in the lower levels. This ensures that bandwidth is not wasted on the client end or on the encoder/server side. The system generates the video stream

to levels 1, 2, 3, 4, 5, respectively. Thus, the total data size of the sample P frame for all five levels is 4 KB.

A diagram illustrating the five levels of video data that make up a B frame as stored in the file format of the present invention is shown in Figure 7. A sample B frame and each of its five levels of data of varying resolution and quality is shown in the Figure. Each level is shown with a corresponding data size. The data size for the levels is 0.15 KB, 0.35 KB, 0.5 KB, 1 KB, 3 KB which correspond to levels 1, 2, 3, 4, 5, respectively. Thus, the total data size of the sample B frame for all five levels is 5.0 KB.

A diagram illustrating a sample group of pictures (GOP) sequence composed of Key, P and B frames making up a video stream is shown in Figure 8. In this example, the video server has determined that level 2 data should be sent for this GOP. Thus, the Key frame 80, B frames 82, 86, 90 and P frames 84, 88, 92 are shown depicting level 2 data and associated data size. The total data size of the GOP is 2.485 KB.

Video Server Process

The video server portion of the video transport system of the present invention will now be described in more detail. The function of the video server is to accept a remote client connection request, retrieve a local or remote stored file and transmit it to the client. Before and during the transmission of the video information, the server appropriately adjusts the rate of data flow from the server to the client. The rate is adjusted beforehand based on initial estimation of the bandwidth of the data channel. In addition, the data rate is adjusted during transmission using a bandwidth measurement method that uses statistical evaluation of the connection between the server and the client. The dynamic adjustment of the data rate by the server functions to allow the client to receive video having a quality that matches the bandwidth capacity of the connection. Further, during the server/client connection, the client can control the transmission of the data by the server, thus performing a video on demand function.

The acknowledge packet sent by the client comprises an identification of the last received packet, its arrival time and a list of any packets missed since the transmission of the previous acknowledge.

A high level diagram illustrating the sender portion of the video server in more detail is shown in Figure 9. The sender 32 comprises a frame selector 100, packet generator 102, packet transmitter 104, rate control unit 106 and receiver 108. In operation, the frame selector functions to accept the full frame video data containing all five levels of data from the receiver and select out of the five levels of data, the level of data appropriate for the connection with a particular client. The choice of what compression level to send is made on a client by client basis. The frame selector used bandwidth information provided by the rate control unit 106 to determine which of the five levels of data to pass to the packet generator. It is important to note that the raw video source data may me compressed into more or less than five levels. A higher number of levels permits a finer tuning of the available bandwidth to the amount of data sent over the connection.

In combination with the estimated bandwidth measurement, the frame selector utilizes a level bandwidth table in determining which level data to select. A different level bandwidth table is associated with each video source file. The level bandwidth table contains an entry for each of the five possible compression levels. Each entry contains the average bandwidth necessary to transmit the data at that level. The frame selector chooses a level having the most information content that the network connection can support using the bandwidth measurements performed by the rate control unit. For example, the level bandwidth table for a sample video source file may be as follows.

Level	Bandwidth Required (Kbps)
5	200
4	100
3	50
2	20
1	10

If, for example, the rate control unit measures the bandwidth of the network connection to be 25 Kbps, the frame selector would pass only level 2 data to the packet generator. Thus, the output of the frame selector would comprise a sequence of video frames wherein each video frame contains data from only one of the video compression levels (level 2 data in this example).

It is important to note for the very first video frame or packet that is to sent to the client, no bandwidth measurement is available. This is because, the bandwidth measurement method, as described in more detail below, utilizes transmitted packets to determine the bandwidth of the channel. Thus, before the first packet is sent, a different mechanisms is used to initially determine the bandwidth of the channel. In its request to open a video source, the video client transmits to the server the bandwidth of the connection the last time the client was connected to a server. This mechanism is based on the assumption that the previous connection a client had with a server is similar to the present connection. In the case where a computer is attached to TCP/IP networks via two ways, e.g., dial up modem and high speed LAN, this mechanism does not provide an accurate initial bandwidth estimate.

The packet generator 102 functions to receive the frames having video data from a particular compression level and encapsulate them into packets for transmission over the network. The assembled packets are output to the packet transmitter 104 which is responsible for delivery of the packets over the network. In addition, to preparing packets from the frames received, the packet generator functions to determine which (if any) frames to skip. Depending on the measured bandwidth of the channel, the packet generator may skip frames in order to reduce the transmitted bit rate. This occurs when the bandwidth of the network connection cannot support transmission of every Key, P and B frame. The method of choosing which frames to select is described in more detail hereinbelow.

The packet generator does not send packets to the packet transmitter 104 until requested to do so by the packet transmitter. The delivery of the packets onto the network is controlled by the rate control unit 106. The rate control unit

keeps track of the amount of video information in terms of time that is queued for display at the client. In addition, the video frames from the video source are time stamped for synchronization purposes. The rate control unit uses acknowledges received by the client via the acknowledgment receiver 108 to determine the next packet transmission time. Once the packet transmitter is notified to send the next packet of data, it requests a packet from the packet generator.

Notification of acknowledges or ACKs received by the receiver 108 are also input to the packet transmitter in order to assure proper receipt by the client. In addition, the packet transmitter maintains a buffer of packets transmitted to the client. In the event the video server determines to resend a packet, the packet transmitter retrieves the packet from the buffer. Once receipt of a packet is acknowledged by the client, the packet is deleted from the buffer and the buffer space is freed up.

Network Bandwidth Measurement Process

The bandwidth measurement method as executed by the rate control unit 106 in the sender will now be described in more detail. The bandwidth measurement method actually comprises two separate phases. The first phase being a scanned phase and the second being a fixed phase. In general, the bandwidth measurement method operates by transmitting packets through the network connection and measuring the rate of reception of the packets at the client. A graph illustrating the receiver bit rate versus the number of bytes online is shown in Figure 10. The number of bytes transmitted into the network pipe is increased slowly until a point is reached where bytes are not received any quicker at the client. The term bytes on line means the number of bytes or packets that have been transmitted by the server or the sender but not yet received by the client. During this scan phase portion of the bandwidth measurement method, the 'immediate' flag is set 'on' for each packet sent by the sender. This causes the client to send an acknowledge packet for every packet received. Thus, the sender should receive an acknowledge packet for every packet transmitted to the client. As shown in Figure 10 as the number of packets or bytes online increases, a point is reached where the client does not receive packets any faster. The

corresponding receive rate at this point can be modeled as an estimate of the bandwidth of the network channel.

The scan phase portion of the bandwidth measurement method will now be described in more detail. A high level flow diagram illustrating the scan phase of the bandwidth measurement method of the present invention is shown in Figure 11. As stated previously, the immediate flag is set 'on' for all packets transmitted by the sender during the scan phase of the bandwidth measurement method. This forces the client to immediately send an acknowledge packet for every packet received over the channel. In addition, an acknowledge packet is also sent if the last received packet has a sequence number greater than the sequence number of the last received packet. In this case, a packet loss event has occurred. Also, an acknowledge packet is sent if the previous acknowledge was sent more than an predefined time out period ago. For example, if the time out period is 3 seconds, an acknowledge is sent if the last packet was received more than 3 second ago.

The acknowledge packet sent by the client contains an identification of the last received packet, it's arrival time and a list of any packets missed since the transmission of the previous acknowledge. Initially, the recommended bytes online (RecommendedBytesOnline) is set equal to the size of the packet (PacketSize) (step 110). In the next step, a single packet is sent by the sender to the client (step 112). The current number of bytes online (BytesOnline) is then calculated (step 114). The number of BytesOnline can be calculated since the sender has knowledge of each packet that is placed into the network pipe in addition to having knowledge of each acknowledgment received from the client. Thus at any one time the sender is aware of outstanding packets still in the network pipe. Next, the number of bytes online is compared to the recommended bytes online (step 116). The number of bytes online can be calculated using the sequence number of the last packet that was sent, as known by the sender, subtracted from the sequence number of the last packet acknowledged. Both these entities are known by the sender and thus the number of bytes online can be calculated. If the number of bytes online are less than the recommended

It is then determined whether the receiving bandwidth has leveled off (step 130). With reference to Figure 10, in this step, it is checked whether the number of bytes online has begun to level off as shown in the right most portion of the curve in the Figure. The leveling off of the receive bandwidth is detected by comparing the current receiving bandwidth to the average of the last five values of the receiving bandwidth. If the latest value of the receiving bandwidth is within 5% of the average then the receiving bandwidth is considered to have leveled off. Consequently, the bandwidth of the network connection is estimated to be the value of the last received bandwidth. If the receiving bandwidth has not leveled off, i.e., within 5% of the average of the previous five measurements, then the recommended bytes online (RecommendedBytesOnline) is incremented by the packet size (step 132). Control then returns to step 112 and an additional packet is placed into the network pipe.

If the receiving bandwidth is found not to have leveled off it means the number of bytes online corresponds to the linear portion of the curve in Figure 10. Thus, the maximum bandwidth of the network pipe has not been reached and additional packets can be pumped into the network channel. If the receiving bandwidth has been found to have leveled off the recommended bandwidth (RecommendedBW) is set equal to the current value of the receiving bandwidth (ReceivingBW) (step 134). The recommended bandwidth value is utilized by the rate control unit as an initial estimate of the bandwidth of the network connection.

The scan phase portion of the bandwidth measurement method is used initially as a relatively crude estimate of the bandwidth of the network channel. During steady state operation of the sender portion of the video server a fixed phase bandwidth measurement method is utilized to better fine tune and track changes in the bandwidth of the network channel. A high level flow diagram illustrating the fixed phase method of the bandwidth measurement portion of the present invention is shown in Figure 12. During the fixed phase of the bandwidth measurement method the immediate flag is set to 'off' in each packet sent by the sender. The first step is to set a variable representing the time to send (TimeToSend) equal to the current time, i.e., now (step 140). Next, it is checked

CLAIMS

1. A method of transporting video over a network channel, comprising the steps of:

   compressing a raw video source into a plurality of frames, each frame comprising a plurality of levels, each level corresponding to a particular degree of compression;

   estimating the bandwidth of the network channel;

   selecting one of said plurality of levels of each frame to transmit over the network channel in accordance with said bandwidth estimate whereby the level selected optimizes the use of the bandwidth of the network channel; and

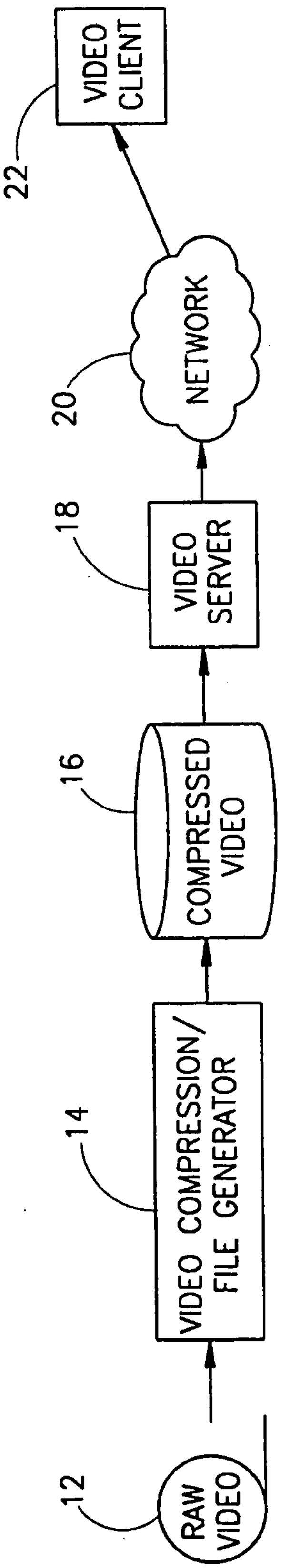
   sending said selected level of each frame over the network channel.

2. The method according to claim 1, wherein said step of compressing comprises the step of compressing the raw video source into a plurality of different types of frames, each frame type containing different amount of video content information, said plurality of different types of frames grouped so as to form a video stream consisting of a plurality of group of pictures (GOP) sequences.

3. The method according to claim 1, wherein said step of compressing comprises the step of compressing the raw video source into Key, P and B type frames, said Key, P and B frames generated so as to form a video stream consisting of a plurality of group of pictures (GOP) sequences.

4. A method of transporting video from a video server to a video client over a network channel, comprising the steps of:

   compressing data from a raw video source so as to generate a plurality of frames, each frame being of a particular frame type, each frame type containing a particular amount of video content information, each frame comprising a plurality of levels, each level corresponding to a particular degree of compression;


RAW VIDEO
12
VIDEO COMPRESSION/ FILE GENERATOR
14
COMPRESSED VIDEO
16
VIDEO SERVER
18
NETWORK
20
VIDEO CLIENT
22

10

FIG.1

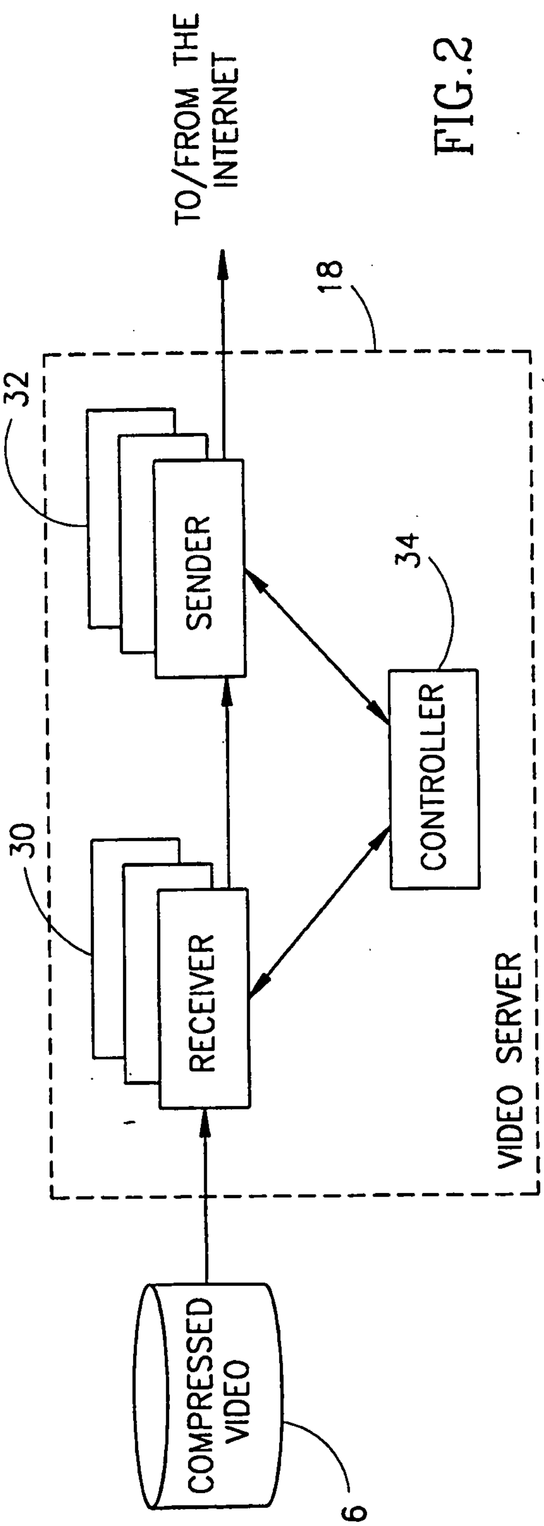

COMPRESSED VIDEO
16
RECEIVER
30
SENDER
32
CONTROLLER
34
VIDEO SERVER
18
TO/FROM THE INTERNET

FIG.2